Fall 2014
SEI Research Review
Malware Distribution Networks

Software Engineering Institute
Carnegie Mellon University
Pittsburgh, PA 15213

Team: Dr. Jose A. Morales
Dr. William Casey
Aaron Volkmann

(presenting)

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
01 OCT 2014	**N/A**	

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Fall 2014 SEI Research Review Malware Distribution Networks	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
Morales /Jose A.	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NUMBER
Software Engineering Institute Carnegie Mellon University Pittsburgh, PA 15213	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited.

13. SUPPLEMENTARY NOTES
The original document contains color images.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**SAR**	**32**	

This material is based upon work funded and supported by the Department of Defense under Contract No. FA8721-05-C-0003 with Carnegie Mellon University for the operation of the Software Engineering Institute, a federally funded research and development center.

DM-0001740

Project Description

Create an approach to graph the topological structure of a domain name based malware distribution network (MDN) by leveraging search engine data that facilitates the identification and attribution of persistent sub-networks and highly trafficked individual domains

Expected Outcomes

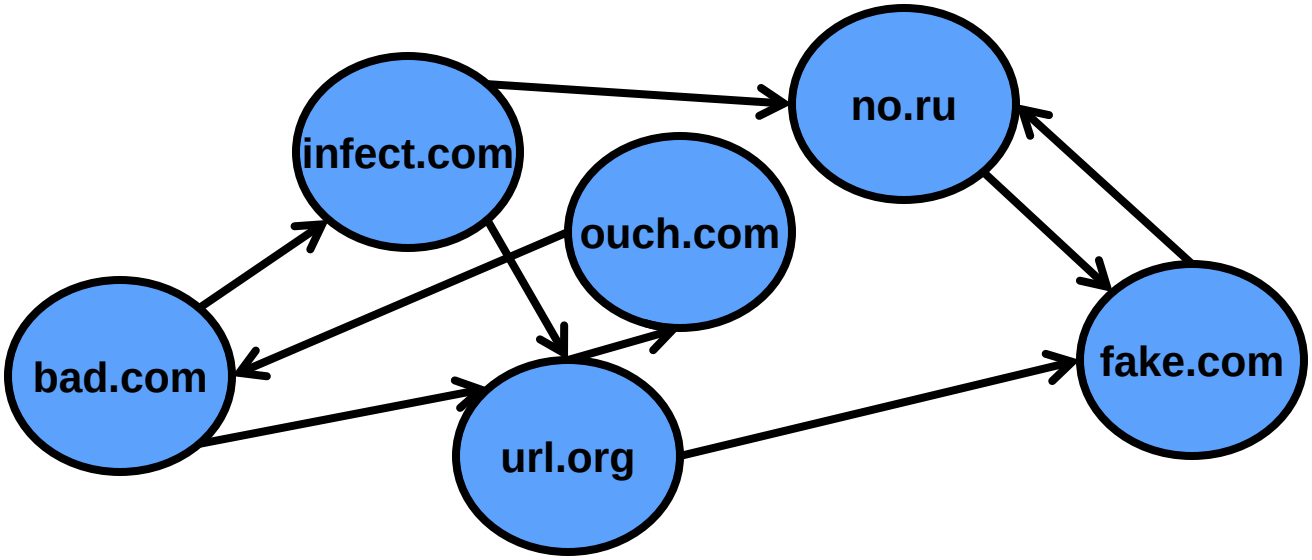
- Identify domains' roles in distribution
- Identify key domains and persistent sub-networks
- Determine MDN structural robustness
- Perform trend analysis to predict future cyber attacks
- Correlate data trends to known offensive/defensive cyber events

Impact for the DoD: Real time tracking of MDNs facilitates identifying early warning indicators of cyber events including potential threats to DoD cyber assets. MDN analysis allows attribution to geographic locations of key malicious resources.

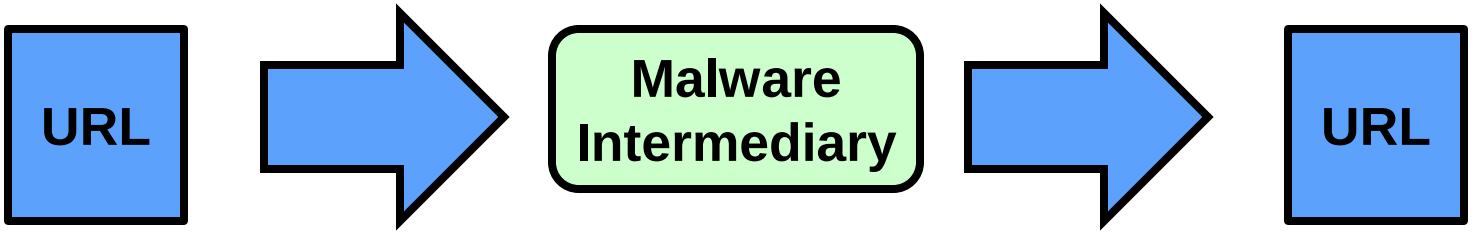
Malware Distribution Network (MDN)

An MDN is an active network of interconnected servers running as a backend to facilitate malware distribution, malicious attacks and other nefarious acts.

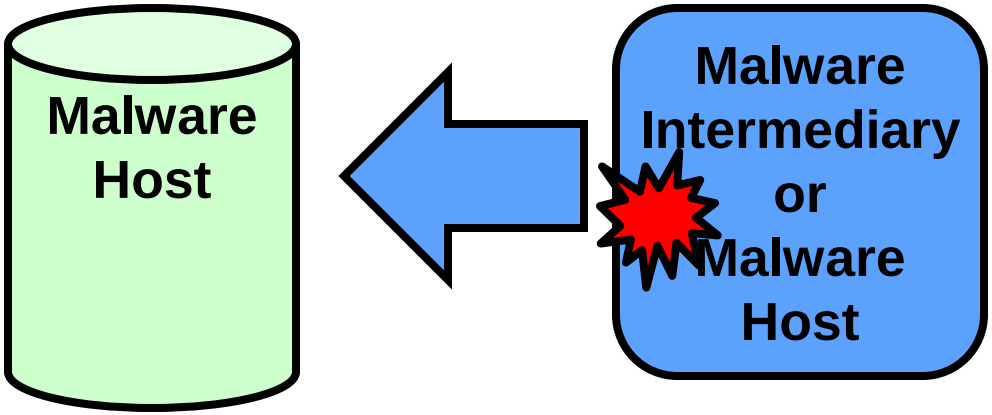
The topological structure of an MDN is represented with a directed graph. Each node is a malicious domain and each edge represents a direct connection between 2 nodes.



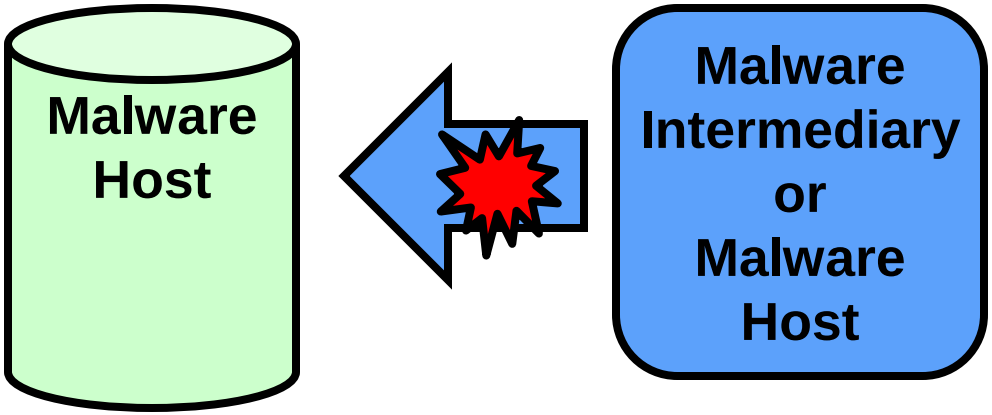
Node Types: Malware Intermediary (MI)



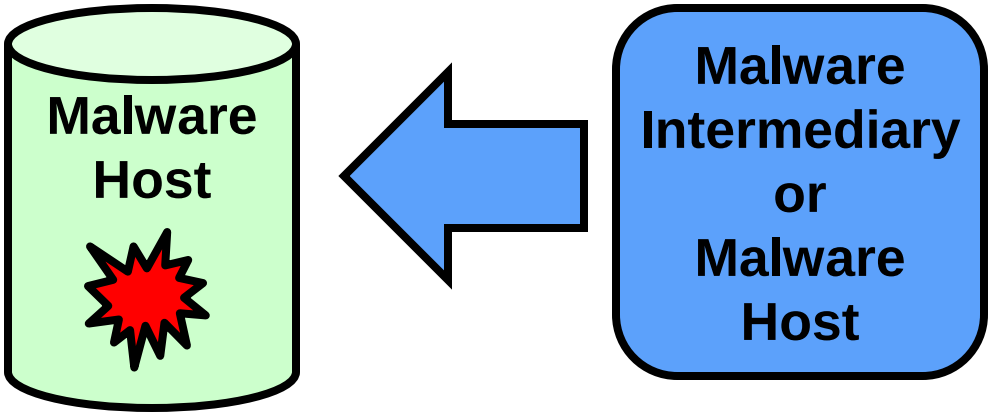
Node Types: Malware Host (MH)



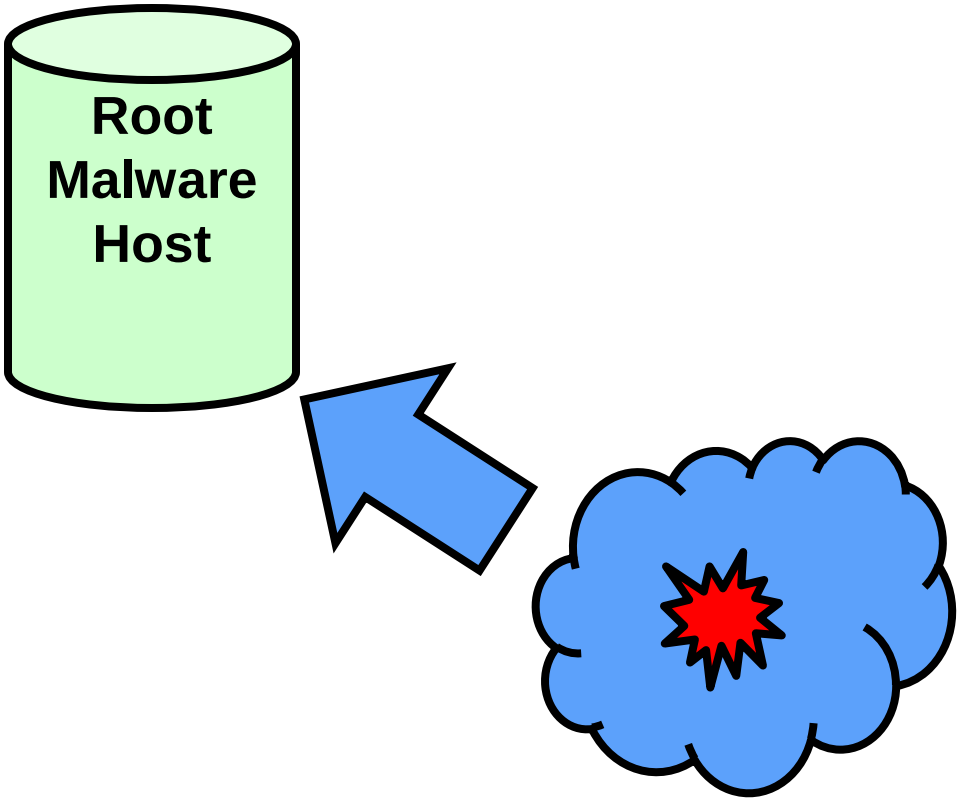
Node Types: Malware Host (MH)



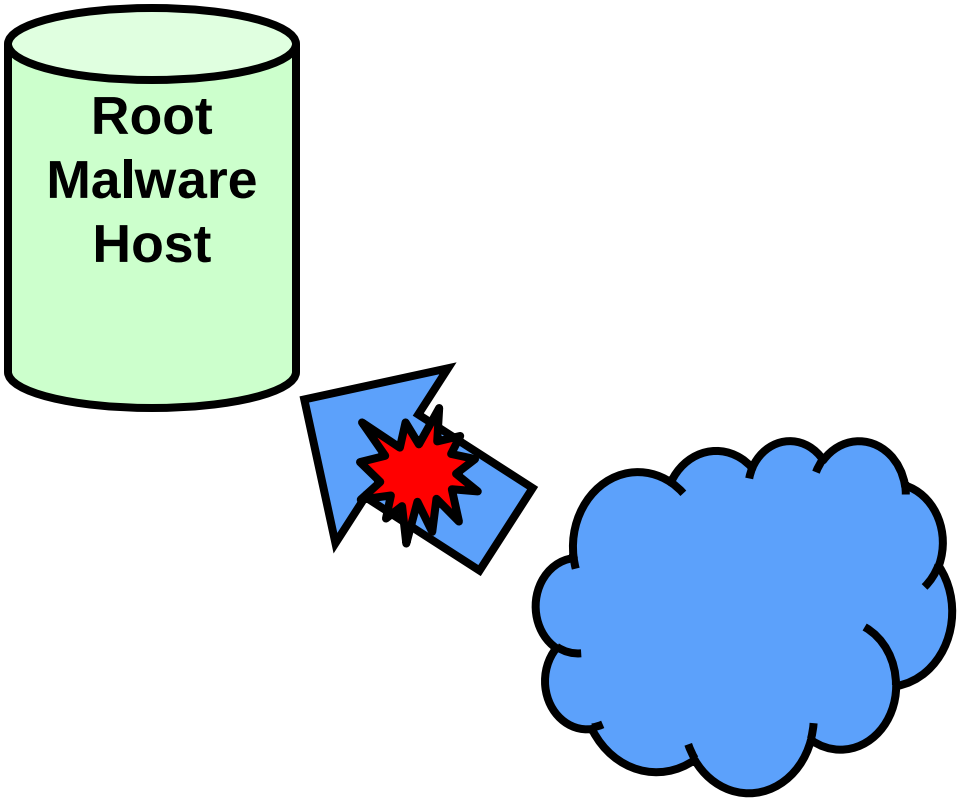
Node Types: Malware Host (MH)



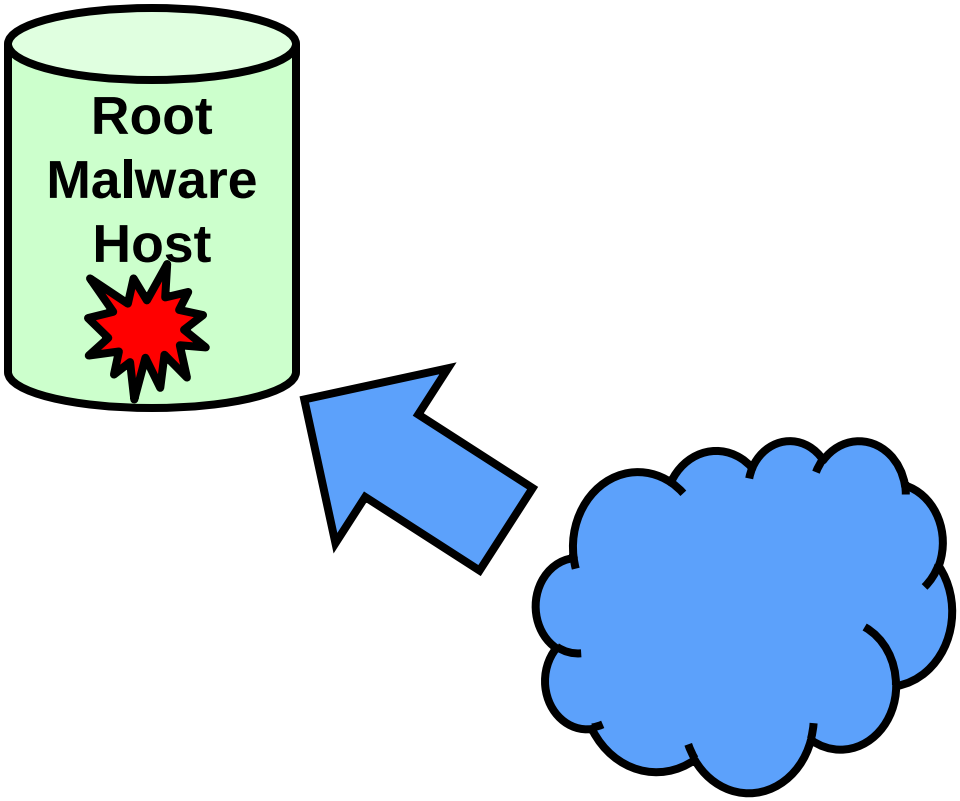
Node Types: Root Malware Host (RMH)



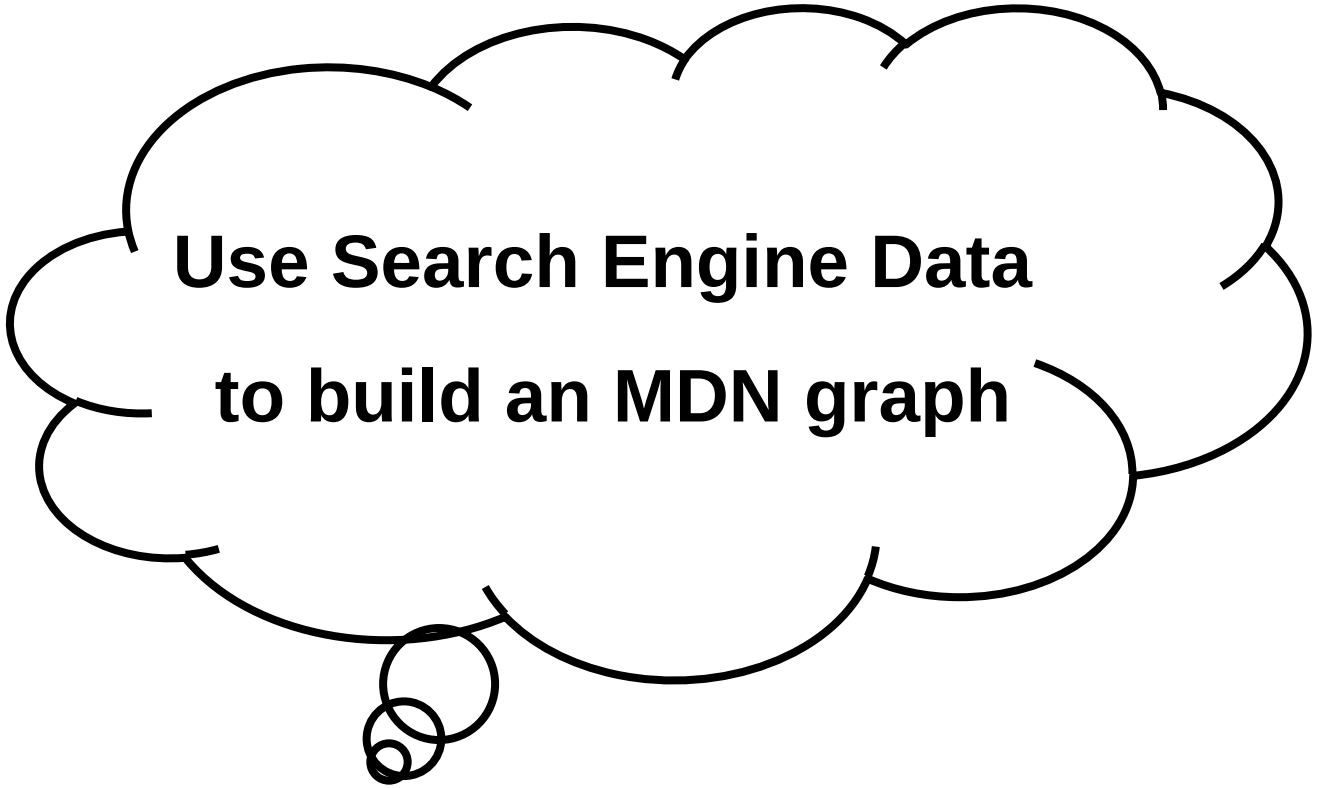
Node Types: Root Malware Host (RMH)



Node Types: Root Malware Host (RMH)



Data Collection



Data Collection

Bing Link From Domain

www.sample.org

www.link.com
www.example.com
free.net
www.trades.com
my.screensaver.co.uk
www.phone.org
www.screensaver.ru

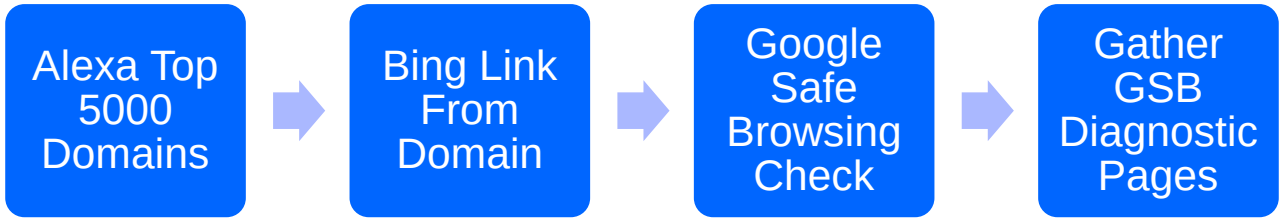
~42 results per domain

Data Collection

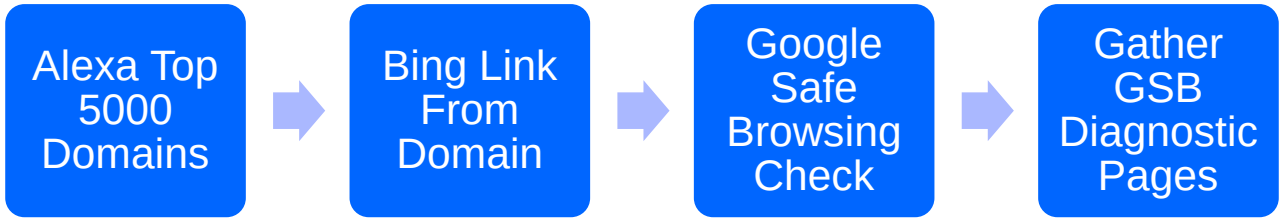
Google Safe Browsing (GSB)

- ✔ www.link.com
- ✔ www.example.com
- ✖ trade.free.net
- ✔ www.trades.com
- ✔ my.screensaver.co.uk
- ✖ www.phone.org
- ✖ www.screensaver.ru

Data Collection



Data Collection



– We collect 3 times a day using 1 Windows & 5 Linux Systems

Graph Creation

Safe Browsing
Diagnostic page for overthehedgemovie.com

Advisory provided by Google

What is the current listing status for overthehedgemovie.com?

Site is listed as suspicious - visiting this web site may harm your computer.

Part of this site was listed for suspicious activity 23 time(s) over the past 90 days.

What happened when Google visited this site?

Of the 20 pages we tested on the site over the past 90 days, 19 page(s) resulted in malicious software being downloaded and installed without user consent. The last time Google visited this site was on 2013-02-20, and the last time suspicious content was found on this site was on 2013-02-20.

Malicious software includes 28 trojan(s).

Malicious software is hosted on 1 domain(s), including hostads.cn/.

This site was hosted on 2 network(s) including AS22822 (LLNW), AS36213 (DWASKG).

Has this site acted as an intermediary resulting in further distribution of malware?

Over the past 90 days, overthehedgemovie.com appeared to function as an intermediary for the infection of 6 site(s) including pebcak.de/, yaneznal.ru/, visuellerorgasmus.de/.

Has this site hosted malware?

Yes, this site has hosted malicious software over the past 90 days. It infected 6 domain(s), including pebcak.de/, yaneznal.ru/, visuellerorgasmus.de/.

How did this happen?

In some cases, third parties can add malicious code to legitimate sites, which would cause us to show the warning message.

Next steps:

- Return to the previous page.
- If you are the owner of this web site, you can request a review of your site using Google Webmaster Tools. More information about the review process is available in Google's Webmaster Help Center.

Updated 5 hours ago

Findings - Domains and Connectivity 01

607 collections from Oct 2012 – Aug 2014

Average Graph has **42,571** nodes, **52,046** edges

Unique domain count overall: **224,282**
Daily max: **56,126** min: **27,772**
Per collection max: **55,632** min: **21,720**

Findings - Domains and Connectivity

Most connected super nodes overall:

1. vk.com **1389**
2. bit.ly **570**
3. amazingonlykeys.com **384**
4. t.co **356**
5. reference.com **294**
6. search.com.vn **289**

Average number of occurrences of each node type per collection:

RMH: **8194** MH: **97** MI: **7556** MH+MI: **5394**

Findings - Domains and Connectivity

Total unique top level domains: **253**

Top 5 most occurring TLDs:

Findings - Domains and Connectivity

Total unique top level domains: **253**

Top 5 most occurring TLDs:

5. de 9,374

Findings - Domains and Connectivity

Total unique top level domains: **253**

Top 5 most occurring TLDs:

5. de 9,374
4. org 9,549

Findings - Domains and Connectivity

Total unique top level domains: **253**

Top 5 most occurring TLDs:

5. de 9,374
4. org 9,549
3. net 13,202

Findings - Domains and Connectivity

Total unique top level domains: **253**

Top 5 most occurring TLDs:

5. de 9,374
4. org 9,549
3. net 13,202
2. ru 17,006

Findings - Domains and Connectivity

Total unique top level domains: **253**

Top 5 most occurring TLDs:

5. de 9,374
4. org 9,549
3. net 13,202
2. ru 17,006
1. com 88,552

Findings - Domains and Connectivity

Total unique IP addresses: **56,339**

Findings - Domains and Connectivity

Total unique IP addresses: **56,339**

Top 5 most occurring IP addresses

1. 46.*.*.* 89
2. 213.*.*.*62
3. 195.*.*.*61
4. 80.*.*.* 60
5. 82.*.*.* 57

Findings - Domains and Connectivity

Unique gov domains:

.gov.* 392

.gov 30

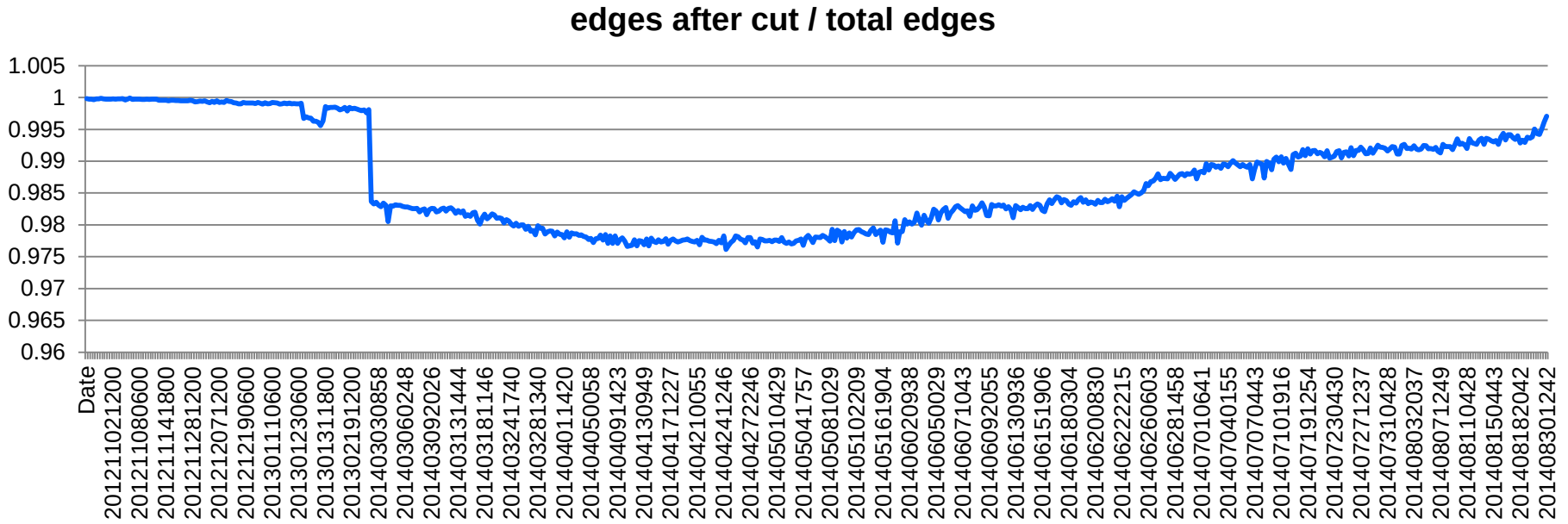
.gov.uk 6

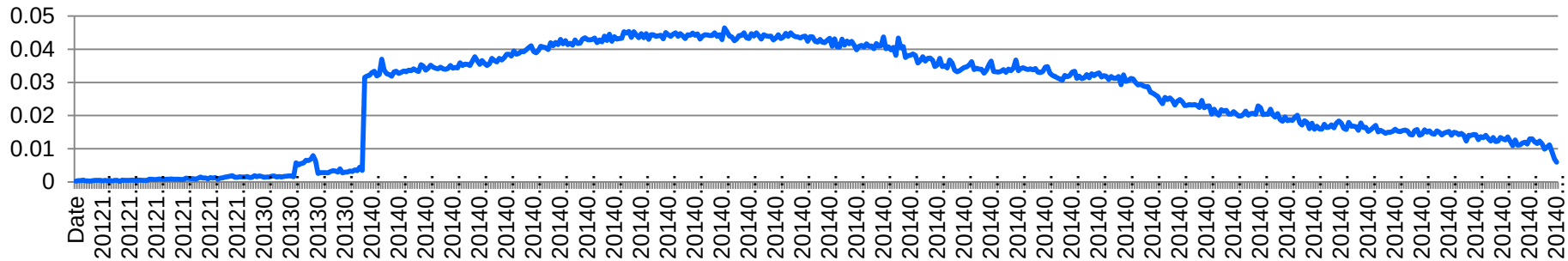
.gov.cn 152

dot gov super nodes:

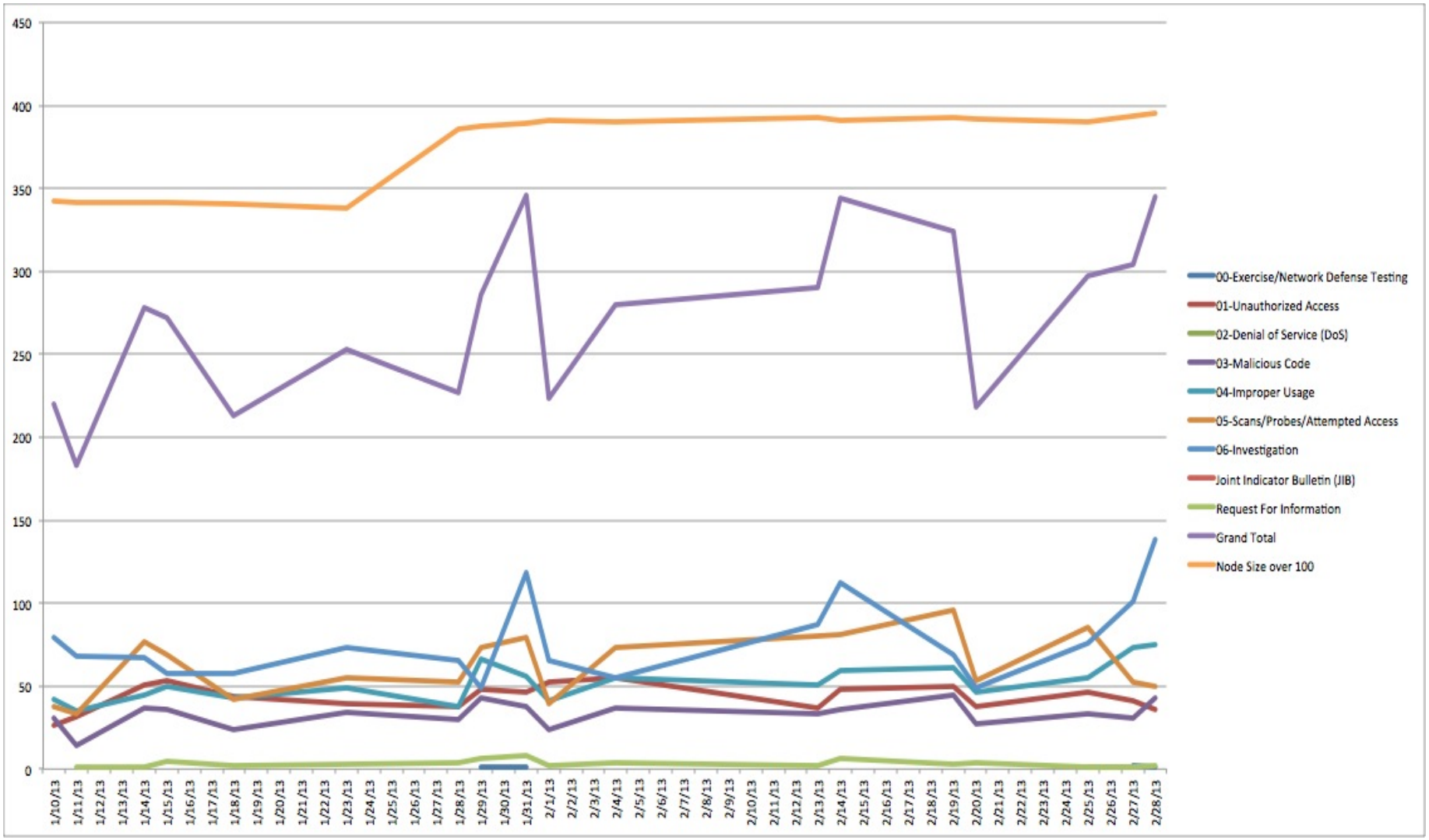
1. 9 edges (1 domain)
2. 8 edges (2 domains)
3. 7 edges (2 domains)
4. 6 edges (3 domains)

Findings – Structural Robustness





Findings – Early Indicators of Cyber Attacks



Conclusions

- MDNs serve as the backend distribution network of malware and malicious cyber events
- A graph can be very large consisting mostly of RMH and MI
- Domains are of all types including .gov
- Structural robustness in minimal, its rather easy to split in subnets
- Evidence suggests MDNs can provide early warning indicators of cyber events

Potential next steps

- Deeper analysis of the collected data
- Attempt the same analysis with other data sets
- Provide early warning indicators to those interested

- We have more detailed data, contact us!

Contact Information

Presenter / Point of Contact

Dr. Jose A. Morales
SEI:CERT
Email: jamorale@sei.cmu.edu

Dr. William Casey
SEI:CERT
Email: wcasey@sei.cmu.edu

Aaron Volkmann
SEI:CERT
Email: amvolkmann@cert.org

U.S. Mail

Software Engineering Institute
Customer Relations
4500 Fifth Avenue
Pittsburgh, PA 15213-2612
USA

Customer Relations

Email: info@sei.cmu.edu
Telephone: +1 412-268-5800
SEI Phone: +1 412-268-5800
SEI Fax: +1 412-268-6257